

Results of Human Papillomavirus DNA Testing with the Hybrid Capture 2 Assay Are Reproducible

Philip E. Castle,^{1*} Attila T. Lorincz,² Iwona Mielzynska-Lohnas,² David R. Scott,³ Andrew G. Glass,⁴
Mark E. Sherman,¹ John E. Schussler,⁵ and Mark Schiffman¹

Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda,¹ Digene Corporation, Gaithersburg,² and Information Management Services, Inc., Silver Spring,⁵ Maryland; Kaiser Permanente Center for Health Research, Portland, Oregon³; and Kaiser Foundation Research Institute, Oakland, California⁴

Received 13 August 2001/Returned for modification 5 November 2001/Accepted 14 December 2001

Reproducibility of the Hybrid Capture 2 Test (HC 2) for human papillomavirus (HPV) DNA detection was evaluated by assaying frozen cervical specimens in 1997 and again in 2001 from 1,775 women with normal cervical cytology. Using a cutoff point of 1.0 pg of HPV DNA/ml between a negative and a positive test result, the result of the kappa test for agreement was 0.72 (a kappa value of >0.60 is considered good agreement). Using cutoff points of 1.0 and 10.0 pg/ml between negative and low positive and between low positive and high positive, respectively, the kappa was 0.68 and the linear-weighted kappa was 0.76. The results of this study indicate that HC 2 testing is reproducible even among cytologically normal women with low test values.

The Hybrid Capture 2 Test (HC 2) (Digene Corporation, Gaithersburg, Md.) (5) is a commercially available assay approved by the Food and Drug Administration for the DNA detection of human papillomavirus (HPV) infections, the primary cause of cervical cancer (1, 6, 11). HC 2 detects a group of 13 cancer-associated HPV types: 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68. Previous studies have demonstrated that HC 2 is a sensitive and specific assay for detecting HPV DNA from cervical specimens (7), perhaps comparable to PCR assays (2), and that HPV DNA detection by HC 2 might be a useful triage for women with a community diagnosis of atypical squamous cells of undetermined significance (9). Understanding the reproducibility of assays based on testing different specimen aliquots or repeat testing over time is important in epidemiologic investigations, especially cohort studies, in which HC 2 is currently being used.

The present analysis compares HPV testing results for 1,775 specimens that were tested in 1997 (test 1) and in 2001 (test 2). The specimens were selected from a collection of 21,145 frozen cervicovaginal lavage specimens obtained from eligible, consenting women between 1989 and 1990 as part of enrollment in a cohort study examining the natural history of HPV and cervical cancer at Kaiser Permanente, Northwest (Portland, Oreg.) (4, 8). Lavages were performed by rinsing the cervical os with 10 ml of sterile physiologic saline using a syringe fitted with an intracatheter extender and then collecting the pooled fluid from the vaginal fornix using the same device (8).

Lavage specimens were refrigerated within 1 h of collection and shipped to a central laboratory for processing. A 1-ml aliquot was removed for HPV DNA testing by PCR as described elsewhere (4, 8). The remaining volume was split into two aliquots of approximately equal volume. Aliquots were either centrifuged at $400 \times g$ for 5 min (4°C) or remained

unprocessed, and the resulting cell pellets or unprocessed lavage aliquots were then frozen at -70°C .

A sample of 1,775 specimens from women in the incidence cohort (no known current or past cervical neoplasia) (4, 8) was chosen for testing. For test 1, cell pellets were used. For test 2, 1,733 of the 1,775 tests were performed on the second aliquot of the lavage stored either as a frozen cell pellet ($n = 803$) or as an unprocessed (liquid) lavage specimen ($n = 930$); 42 specimens were tested on the same lysed aliquot used in 1997 after one additional freeze-thaw cycle.

In test 1, cell pellets were thawed, suspended in 1 ml of specimen transport medium (STM; Digene), and mixed vigorously. Following incubation for 45 min at 65°C with 0.5 ml of the denaturation reagent, a 75- μl aliquot was removed for the HC 2 test, performed according to the package insert. We subsequently discovered that specimens stored for extended periods in the alkaline solution might be unstable for DNA quality. Thus, for long-term storage of DNA, the alkalizing step was excluded from processing of aliquots for test 2. Unprocessed lavage aliquots were first pelleted, and then the freshly pelleted cells were treated similarly to aliquots stored as frozen cell pellets, both of which were resuspended in 1 ml of STM and incubated at ambient temperature for 48 to 72 h. After incubation, the exfoliated cervical cells were lysed and a 100- μl aliquot of the cell lysate (instead of the entire lysed specimen) was denatured using 50 μl of denaturant in a 96-well plate on a plate heater for 45 min at 65°C . This volume was sufficient for two HC 2 tests.

Results were expressed in terms of relative light units of the specimen divided by the mean relative light units of three controls positive for HPV type 16 (HPV-16) DNA (1.0 pg/ml) to estimate the test value in picograms per milliliter. The estimated values in picograms per milliliter were significantly nonnormal even after log transformation ($P < 0.001$). Therefore, to compare the nonnormal continuous in picograms-per-milliliter values, the Spearman (rank) correlation coefficient was determined for all (untransformed) values. Also, because of the preponderance of HPV DNA negative specimens with

* Corresponding author. Mailing address: Division of Cancer Epidemiology and Genetics, National Cancer Institute, 6120 Executive Plaza South, Room 7074, Rockville, MD 20892-7234. Phone: (301) 435-3976. Fax: (301) 402-0916. E-mail: pc95p@nih.gov.

test values below the level of assay accuracy, Spearman correlation coefficients were computed for values above 1.0 pg/ml.

Kappa values (3, 10) (which test for agreement while taking chance agreement into account) were calculated based on a cutoff of 1.0 pg/ml (~5,000 genomic equivalents of HPV DNA) between a negative and a positive test result, which was previously demonstrated to optimize assay performance (7). With the addition of a second cutoff point at 10.0 pg/ml, data were also categorized as negative (<0.995 pg/ml), low positive (1 to 9.995 pg/ml), and high positive (≥ 9.995 pg/ml) and evaluated using both kappa and linearly weighted kappa statistics, which linearly weight disagreement inversely proportional to the number of categories of difference between the two assays (3, 10) (i.e., exact agreement is given a weight of 1, a difference of one category is given a weight of 0.5, and a difference of two categories is given a weight of 0).

All HPV testing was performed independently, with previous test results masked. In test 1, a subset of 200 randomly chosen specimens was assayed in duplicate in a masked fashion for quality control purposes. The overall Spearman coefficient of the picograms-per-milliliter values for the duplicate testing was 0.66; the Spearman coefficient among those specimens with values above the 1.0-pg/ml positive control cutoff point was 0.98. For a negative-versus-positive test based on the 1.0-pg/ml cutoff point, there was 97% exact agreement and a kappa of 0.87 (95% confidence interval [CI] = 0.73 to 1.0) for the

I(29(0.41631.771870 TDO(of)-296.5(the)-294.9(specimens)-294.9(teed,ns)-294.9(there)-294.wthere)-294.1,734he)-294.9(specimens)-291.5(wit

the sndardhe viadition was pg/ml, the rgthe was

pg.so for teor the was pg/or the

was22716.38.2as22716.8(pg/or)2716.5(the)2716.7(sndardhe)2712.7(viadition)2716.[(was22716.305.3as)2716.8(pg/or)2716.6(and)-701.2(the)]T

TABLE 2. Results of HC 2 testing on specimens tested in test 1 (1997) and in test 2 (2001) using a 1.0-pg/ml cutoff between negative and positive and a 10.0-pg/ml cutoff between low positive and high positive

Result by test 1	No. of specimens with indicated result by test 2			Total no. for row
	Negative	Low positive	High positive	
Negative	1,360	63	8	1,431
Low positive	61	66	13	140
High positive	10	16	137	163
Total	1,431	145	158	1,734

comparison of test 1 and test 2. For the results of the test with negative, low-positive, and high-positive cutoff points (Table 2), kappa was 0.68 (95% CI = 0.63 to 0.72) and the linearly weighted kappa was 0.76 (95% CI = 0.73 to 0.80). There were no differences in the reproducibility of HC 2 when the results were stratified by storage conditions (lysate, frozen pellet, or unprocessed lavage) of specimens used in test 2, and therefore only the combined data are reported.

These results suggest that HC 2 testing is reproducible and robust with a 4-year interval between tests. Although there were apparent slight decreases in the average test values (e.g., mean and range) in test 2 compared to test 1, these differences were nonsignificant. Furthermore, the discordance between the tests was symmetrical (McNemar test, $P = 1.0$, for dichotomous outcome; symmetry χ^2 test, $P = 0.90$, for trichotomous outcome), suggesting that there was no bias at one time point versus the other. Of the 142 discordant tests (of 1,734 tests, or 8%), 124 were negative in one test and gave a low-positive result in the other. Therefore, most disagreements were not extreme, as reflected by the weighted kappa being greater than the unweighted statistic. Given that these women were mostly without disease and therefore had a low average viral load, the crude method of dividing of the split specimen employed in this study into aliquots (which likely resulted in variability in cell numbers between pellets), the different specimen preparation used in test 2, and the different cell processing methods for the different HC 2 testing periods, it is perhaps not surprising that a small percentage of tests fluctuated between a negative and a low-positive test result. For 92 of 124 (74%) borderline specimens, HPV DNA testing was also performed by PCR with L1 consensus primers MY09/11 (4); 68% of these specimens were HPV DNA negative by PCR for the same HPV types tested for by HC 2, suggesting that most of these specimens might truly be HPV DNA negative or have very low viral loads.

In summary, HC 2 demonstrated good reproducibility de-

spite testing among women who often had viral loads near the minimum threshold of accurate detection. Our findings suggest that HC 2 can be used on archival specimens for either epidemiologic studies or for retrospective testing in the clinical management of equivocal cytology such as atypical squamous cells of undetermined significance (9), where higher viral loads in these cervical specimens may result in even better assay performance.

We acknowledge the excellent technical support from Julie Buckland in the data analysis group at Information Management Services, from Robert Hallenberg in the HPV testing laboratory at Digene, and from Brenda Rush and Patti Lawler at Kaiser Permanente. We thank the physicians and nurse practitioners at Kaiser Permanente who expertly collected the specimens used in this research in the course of their clinical duties.

A. Lorincz is a Senior Vice President and the Chief Scientific Officer of Digene Corporation. I. Mielzynska-Lohnas is Director, Women's Health, Research, and Development, at Digene Corporation.

REFERENCES

1. Bosch, F. X., M. M. Manos, N. Munoz, M. Sherman, A. M. Jansen, J. Peto, M. H. Schiffman, V. Moreno, R. Kurman, K. V. Shah, et al. 1995. Prevalence of human papillomavirus in cervical cancer: a worldwide perspective. *J. Natl. Cancer Inst.* 87:796-802.
2. Bozzetti, M. C., B. Nonnenmacher, I. Mielzynska, L. L. Villa, A. T. Lorincz, V. Breitenbach, and J. C. Prolla. 2000. Comparison between hybrid capture II and polymerase chain reaction results among women at low risk for cervical cancer. *Ann. Epidemiol.* 10:466.
3. Landis, J. R., and G. G. Koch. 1977. The measurement of observer agreement for categorical data. *Biometrics* 33:159-174.
4. Liaw, K. L., A. G. Glass, M. M. Manos, C. E. Greer, D. R. Scott, M. Sherman, R. D. Burk, R. J. Kurman, S. Wacholder, B. B. Rush, D. M. Cadell, P. Lawler, D. Tabor, and M. Schiffman. 1999. Detection of human papillomavirus DNA in cytologically normal women and subsequent cervical squamous intraepithelial lesions. *J. Natl. Cancer Inst.* 91:954-960.
5. Lorincz, A. 1996. Hybrid Capture method of detection of human papillomavirus DNA in clinical specimens. *Papillomavirus Rep.* 7:1-5.
6. Munoz, N. 2000. Human papillomavirus and cancer: the epidemiological evidence. *J. Clin. Virol.* 19:1-5.
7. Schiffman, M., R. Herrero, A. Hildesheim, M. E. Sherman, M. Bratti, S. Wacholder, M. Alfaro, M. Hutchinson, J. Morales, M. D. Greenberg, and A. T. Lorincz. 2000. HPV DNA testing in cervical cancer screening: results from women in a high-risk province of Costa Rica. *JAMA* 283:87-93.
8. Schiffman, M. H., H. M. Bauer, R. N. Hoover, A. G. Glass, D. M. Cadell, B. B. Rush, D. R. Scott, M. E. Sherman, R. J. Kurman, S. Wacholder, C. K. Stanton, and M. M. Manos. 1993. Epidemiologic evidence showing that human papillomavirus infection causes most cervical intraepithelial neoplasia. *J. Natl. Cancer Inst.* 85:958-964.
9. Solomon, D., M. Schiffman, R. Tarone, and the ALTS Study Group. 2001. Comparison of three management strategies for patients with atypical squamous cells of undetermined significance: baseline results from a randomized trial. *J. Natl. Cancer Inst.* 93:293-299.
10. Stoler, M. H., M. Schiffman, and the Atypical Squamous Cells of Undetermined Significance-Low-Grade Squamous Intraepithelial Lesion Triage Study (ALTS) Group. 2001. Interobserver reproducibility of cervical cytologic and histologic interpretations: realistic estimates from the ASCUS-LSIL Triage Study. *JAMA* 285:1500-1505.
11. Walboomers, J. M., M. V. Jacobs, M. M. Manos, F. X. Bosch, J. A. Kummer, K. V. Shah, P. J. Snijders, J. Peto, C. J. Meijer, and N. Munoz. 1999. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J. Pathol.* 189:12-19.